Korean J. Chem. Eng., 36(10), 1637-1647 (2019)
DOI: 10.1007/s11814-019-0330-y

pISSN: 0256-1115
eISSN: 1975-7220

Modeling of the solubility of H_2S in [bmim][PF_6] by molecular dynamics simulation, GA-ANFIS and empirical approaches

Amir Dashti*,†, Farshid Zargari, Hossein Riasat Harami***, Amir H. Mohammadi****,*****,†, and Zahra Nikfarjam******

*Young Researchers and Elites Club, Science and Research Branch, Islamic Azad University, Tehran, Iran
**Department of Chemistry, Faculty of Science, University of Sistan and Baluchestan (USB), Zahedan, Iran
***Department of Chemical Engineering, University of Kashan, Kashan, Iran
****Institut de Recherche en Genie Chimique et Petrolier (IRGCP), Paris Cedex, France
*****Discipline of Chemical Engineering, School of Engineering, University of KwaZulu-Natal, Howard College Campus, King George V Avenue, Durban 4041, South Africa
******Department of Molecular and Supramolecular Modelling, Chemistry and Chemical Engineering Research Center of Iran, Tehran, Iran
(*Received 18 February 2019 • accepted 27 June 2019*)

Abstract−Predicting the solubility of acid gases in ionic liquids (ILs), has lately appeared as advantageous for natural gas purifying, which is equipped by powerful models considering technical and economic aspects. Important issue in the assessment of ILs for potential utilization in gas sweetening process is estimating the H_2S solubility at various temperatures and pressures Experimental measurements are costly and take considerable time and effort. As a result, proposing methods for predicting the behavior of this system over a wide range of conditions is vital. In this regard, molecular dynamics simulation (MD) technique as well as artificial intelligence knowledge of hybrid genetic algorithm-adaptive neuro fuzzy inference system (GA-ANFIS) and an empirical polynomial regression (PR) model were employed to estimate the solubility of H_2S in [bmim][PF_6] IL. Pressure and temperature are considered as the independent input variables and H_2S solubility as the dependent output variable. The results of this study reveal that the simple fourth-order PR model and GA-ANFIS have the highest accuracy. As a result of the simplicity and accuracy of PR model, it can be used without any prior knowledge about MD and artificial intelligence (AI). According to the accuracy and precision of model proved by the obtained result, the solubility of H_2S in ILs has been estimated. The results show that the PR method is more trustworthy than other models.

Keywords: H_2S, Ionic Liquid, GA-ANFIS, Regression, Molecular Simulation

INTRODUCTION

Eliminating acidic contaminants such as H_2S and CO_2 gases from oil and gas streams is imperative and directly relevant to environmental conservation and committing safety criteria as much as economic reasons [1-4]. H_2S toxicity and its emission to the atmosphere are the two most essential reasons to keep it from being airborne. This method is the most commonly used in absorption by Alkanolamine solutions followed with desorption, which is carried out by reducing the pressure and heating the solution. Besides, Disopropanolamine, Diethanolamine, Mono-Ethanolamine and N-methyl-Diethanolamine are the most prominent industrial Alkanolamines for this process. In recent years, ILs have attracted attention, for not only reduce the loss of Alkanolamine, but also deplete the amount of water present in the gas stream. Besides, by utilizing ILs, lower energy will be needed in comparison with high energy demand to regenerate the aqueous Alkanolamines [5,6]. ILs cover a vast range of various new syntheses because of interesting characteristics for wiping out the hazardous and polluting organic solvents [7]. A group of astonishing benefits of ILs compared to other organic solvents can be mentioned as high thermal conductivity, capability of ILs for dissolving many organic, inorganic and organometallic compounds, being highly polar and non-aqueous polar alternatives. Additionally, very low vapor pressure is the most noticeable feature [8-14].

Gas solubility at different temperatures and pressures plays the major role in evaluating the proficiency of ILs for gas sweetening systems [15-17]. There are numerous papers concerning gas solubility, mainly CO_2 in ILs, though the investigational data of H_2S solubility in ILs are limited [18-22]. Subsequently, it is undoubtedly priceless to develop predictive approaches for appraising and predicting the properties of systems over various conditions.

During the last decade, many researchers were advancing simple models to complex ones to approach the prominent model of solubility in ILs. Kamps et al. studied the solubility of CO_2 in [bmim][PF_6] in the temperature range of (293 to 393) K and pressures up to about 9.7 MPa. Solubility values were correlated by

†To whom correspondence should be addressed.
E-mail: amirdashti13681990@gmail.com, amir_h_mohammadi@yahoo.com

means of the extended Henry's law [23]. In another mini review, Shukla et al. considered the superior potential of IL to absorb and capture CO_2 gas molecules; besides, the nature and mechanisms of interactions between IL-CO_2 were investigated, which completely depend on the nature of cation, anion and presence of functional group. Note that the extent of interaction between the components of ILs plays an important role in CO_2 capture [24]. Moreover, Ghazani et al. developed least square support vector machine (LSSVM) algorithm as a suitable model to estimate the solubility of CO_2-rich gaseous mixtures including H_2S, SO_2, CH_4, and N_2O in different ILs [25]. The solubility of CO_2 in 20 different ILs over a vast range of pressure (0.25-100.12 MPa) and temperature (278.15-450.49 K) was predicted by Mesbah et al. [26]. They used a robust MLP-NN method with the R^2 of 0.9987, MSE of 0.6293 and AARD% of 1.84 which proved the model was accurate in estimating experimental values. By considering the researchers' published papers, it becomes clear that, the most repeatedly used simple cubic equations of state (EOSs) which have been utilized by researchers are Peng-Robinson (PR) [27] and Soave-Redlich-Kwong (SRK) [28]. In addition, in these approaches instead of linking an anion and a cation, the IL was demonstrated as a complete molecule comprising cohesive energy and certain volume [29]. Other results of attempting in ILs modeling are the activity coefficient models and group contribution methods. In another aspect, several excess Gibbs energy models, such as the Non-random two-liquid (NRTL) model [30], Wilson's equation [31], UNIQUAC [32] and the group contribution method of UNIFAC [33] have been scrutinized in several articles. Since ILs have several exclusive features like insignificant vapor pressures and long alkyl chains on the cations in which the polymers have the same characteristic, a broad range of investigations have endeavored to model ILs by lattice models. The statistical associating fluid theory (SAFT EoS) is an effective solubility model for ILs where its modifications, like PC-SAFT [34], tPC-PSAFT [35] and SAFT-γ [36], have been provided as well [29]. It is noteworthy that these sorts of complicated thermodynamic models which consist of complex calculations need various mathematical derivations and innovative programming. Regardless of the particulars or exceptions, there is no general approach for all the mentioned methods. Also, they are limited to the particular systems in which they have been modulated [37,38]. Hence, the capability of estimating the phase behavior in such systems for different sorts of ILs needs applying general and robust approaches which cover various conditions considering the supreme importance.

MD simulation is a tremendous method to predict the gas solubility. This method was used for the bulk to obtain their equilibrium and non-equilibrium properties such as thermodynamic and transport properties of wide varieties of materials. However, the acute bottleneck in using MD to obtain the accurate and reliable properties is determining the energy parameters of the interaction of every single atom in the simulated system, which is called force field [39,40]. Many methods have been developed to calculate the solubility of gases in various materials based on molecular simulations. Some of the most important and at the same time, most commonly used methods are the Monte Carlo (MC) methods. Other methods were developed on the basis of small equilibrium cycles in the form of *NPT* and *NVT* ensembles using MD simulation [41]. To achieve the solubility of benzene in dimethyl-imidazolium chloride and dimethylimidazolium hexafluorophosphate ILs, Hanke et al. [42] tried to use MD simulations considering the benzene with and without charges. One of the solutions to get the solubility of carbon dioxide in [C_4mim][PF_6] with united-atom (UA-FF) and all-atom (AA-FF) force fields [43,44] is to use *NPT* MC simulations, which was carried out by Shah and Maginn [45] who suggested this method and applied Widom's test particle insertion [46]. By investigating the H_2S solubility in some ILs, Pomelli et al. [47] indicated that the acid gas does not have a reaction with the IL. By the first principle calculations of the interaction energy of H_2S with [C_4mim][Cl], [BF_4], [TfO], [Tf_2N] and [PF_6], they found that hydrogen sulfide was highly solvable in [C_4mim][Cl], whereas the solubility of H_2S in [C_4mim][PF_6] was very much lower. Jalili et al. [48] noticed that H_2S is more soluble compared to CO_2 by studying the solubility of CO_2, H_2S and their mixtures in [C_8mim][Tf_2N]. They conducted *ab initio* calculations and explained that the energetic interactions of H_2S with cation part of IL and the anion part [Tf_2N] were higher as compared to the CO_2-IL and CO_2-[Tf_2N]. The authors mentioned that the main reason for solubility is the strength of gas-anion interactions. The amount of available information for the H_2S-IL system, whether theoretical or experimental, is limited and available for different ILs like [C_4mim][PF_6], [C_4mim][BF_4], [Cnmim][Tf_2N] and [C_4mim][Cl] [15,49], while the amount of existing information for CO_2 is considerable and proceeding [23, 50-52].

Recently, the capability of analyzing and modeling enigmatic and challenging issues has lead to several researches implemented widely in science and engineering by some soft computing methods, like support vector machines (SVMs), GAs, ANFIS and artificial neural networks (ANNs) [53]. The solubility behavior of various gas-liquid systems has been surveyed by several ingenious models recently. Shafiei et al. attempted to calculate the H_2S solubility in eleven different ILs by the use of ANN model, two different optimization algorithms under the name of particle swarm optimization (PSO) and back propagation algorithms were used. According to the obtained data, the accuracy of PSO-ANN was higher than back propagation-ANN [54]. Having considered the capability of least squares support vector machine (LSSVM) equipped with an optimization method, under the name of GA, the H_2S solubility in eleven different ILs was modeled by Ahmadi et al. Additionally, the GA-LSSVM outputs were compared with outputs of SRK [28] and PR [27] EOSs. The obtained results indicated that the optimized GA-LSSVM model demonstrated higher accuracy than SRK and PR EOSs [55]. On the other hand, an -ANN and an ANFIS as well as SRK and PR EOSs were used by Baghban et al. [56] for estimating the CO_2 solubility in fourteen different ILs. It was revealed that the ANN approach was superior to the ANFIS model and the other two EOSs [56].

Current study spotlights the application of new methods covering the prediction of H_2S solubility in [bmim][PF_6]. Moreover, calculation of CO_2 solubility in many commonly investigated ILs has been represented as a unique correlation relying on a multivariable regression method [24].

In this study, for the first time, three various modeling methods, namely GA-ANFIS as an AI method, new emerging technol-

ogy of molecular simulation, and simple yet accurate model of polynomial regression, are proposed and compared with each other to yield a guideline for researchers to utilize the best model in their studies. In addition, the present study aimed to carry out a comprehensive atomistic MD simulation of [bmim][PF_6]. Actual data were gathered from the recent research in the literature. Accuracy and precision of developed models were surveyed via various statistical and graphical approaches.

1. [bmim][PF_6]

Regarding difficulties such as the non-ideal behavior of studied systems, the complicated structure of ILs, high cost processes and also time-consuming measurement, modeling methods have superior priority to estimate the gas solubility compared to carrying out experimental measurement. One of the most common types of ILs is [bmim][PF_6], which comprises two unlike electrical charges, $[bmim]^+$, $[PF_6]^-$; also, it is a colorless, viscous, non-water soluble and hydrophobic IL that its melting point is reported around 265.15 K. [PF_6] anions are approximately 0.6 Å closer to the [bmim] cation and a more complex solvation structure (three layers) inferred than for [BF_4], which is another common type of ILs. Not only this deviation makes the [bmim][PF_6] a superior candidate to attract CO_2 gas molecules, but also forms stronger hydrogen bonding, which leads to more stable material. Besides, significant solubility of CO_2 in [bmim][PF_6] IL is reported in literature [57,58].

THEORY

1. MD Simulation Details

In this study, the sorption module of the Materials Studio was utilized. It is an MC method that is specifically appropriate for adsorption phenomena estimation over defined adsorbents. The criterion for choosing the Materials Studio sorption module is that it has been used by many researches for predicting the absorption of gases in porous materials. It has been used to simulate a pure sorbate (or a mixture of sorbate components) absorbed in a sorbent framework, typically a microporous crystal such as a zeolite [59], metal organic frameworks (MOF) [60], polymers [61], and so on [62-64].

In the present work, BMIM and PF_6 were constructed using the sketch module of material studio. To assign the appropriate partial charge to the atoms, DFT calculation using Gaussian03 was done with B3LYP [65] and basis set 6-311+G* [66]. The electro-static potential (*ESP*) fitting method was used to analyze of partial charge for optimizing structures of anion and cation. The final structures with partial charge labeled for each atom of bmim and PF_6 are shown in Fig. 1.

The bulk IL was constructed by the Amorphous Cell Module in Material Studio, where the initial density with the initial guess was set to 1.4 g/cm^3 which was close to experimental density. Periodic boundary conditions were applied to the constructed cell where it contains 100 cations and 100 anions. An amorphous cell constructed in the current simulation is illustrated in Fig. 2 as a sample. The constructed model was then subjected to the NVT simulation in 300 K for 200 ps to equilibrate the temperature of the system. 200 ps-NPT simulations were applied at 1 atm pressure and 300 K temperature to adjust the pressure of the system. The selected

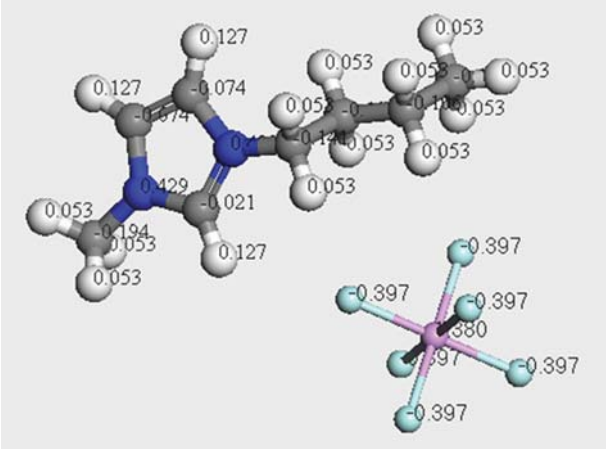

Fig. 1. Structures and partial charges for the cation and anion studied in this work.

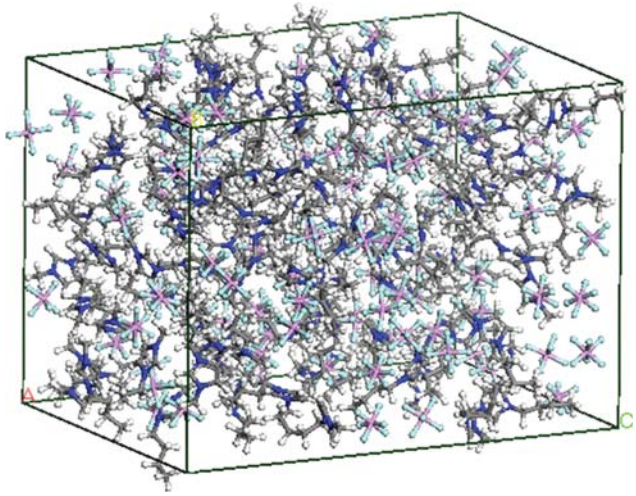

Fig. 2. A sample of unit cell of [bmim][PF_6] built by 'Amorphous Cell' module in Material Studio.

temperature and pressure in these steps were independent of the temperature we chose in the metropolis step, and they were just assigned to relax the initial structure. The NVT run was performed for 200 ps to make sure it was enough to completely relax the system at 300 K. It converged after 50 ps simulation in 300 K (Fig. 3). To relax the density of the initial structure we also performed the NPT run for 200 ps. Fig. 4 shows the changing density of the system versus time. The density of the system converged at about 40-50 ps and was fluctuating about the average amount of 1.475 g/cm^3.

Berendsen barostat and Andersen thermostat were marked to perform an NPT MD simulations [67]. Having used Forcite module, MD simulations were performed. Also, atomic base tab was chosen in the summation method for electrostatic and van der Waals interactions. A 15 Å-repulsive cut off was used for van der Waals interactions. In the forcefield type, COMPASS II was chosen to perform all MD simulation runs. By taking advantage of MC sampling, sorption module of Material Studio was used for an alternative method to calculate the sorption of H_2S into [bmim][PF_6].

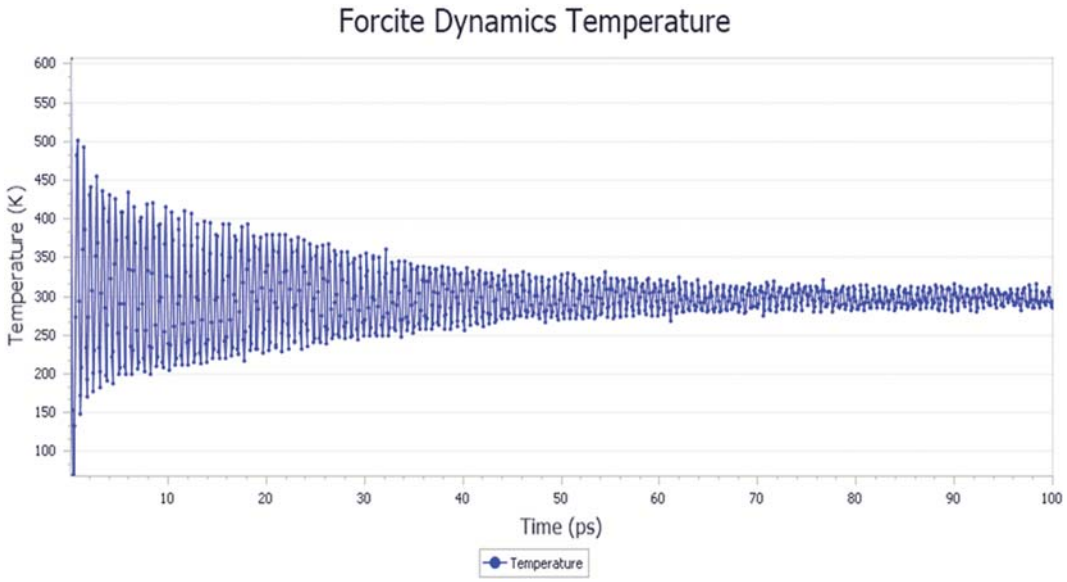

Fig. 3. The temperature of [bmim][PF_6] during NVT simulation.

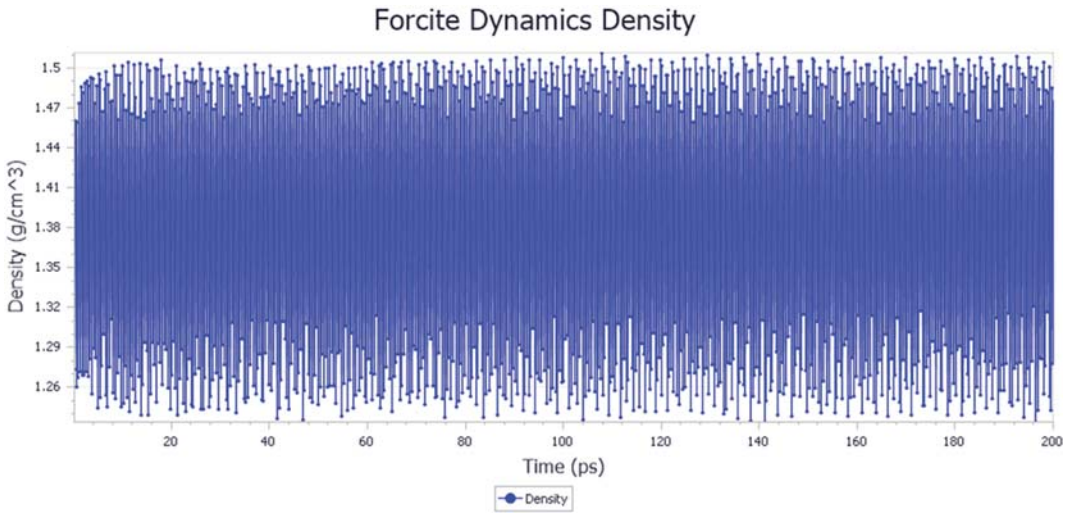

Fig. 4. The density of [bmim][PF_6] during NPT simulation.

2. ANFIS

ANFIS is an AI technique adjusted to map difficult and highly nonlinear systems. Initially, the approach was proposed by Jang [68]. ANFIS not only integrates both ANN and FIS, but also handles complex and nonlinear issues in a single framework. The simple structure of ANFIS includes two inputs and five layers. In this structure, the Takagi-Sugeno fuzzy type is employed as FIS. The FIS, including one output (O) and two inputs (k, j), is assumed in order to define the procedure of ANFIS. The fuzzy rules are reported as follows [69]:

Rule 1:
If x_1 is A_1 and x_2 is B_1 etc., then $f_1=p_1x_1+q_1x_2+\ldots+r_1$;

Rule 2:
If x_2 is A_2 and x_2 is B_2 etc., then $f_1=p_2x_1+q_2x_2+\ldots+r_2$;

The ANFIS architecture consists of a feed-forward network, containing a five-layer structure with different functions. Eqs. (1)-(6) show the function of each layer. The association between the input and output functions of layer 1 including input nodes is described as follows:

$$Q_{1,i}=\mu_{Ai}(m)i=1, 2 \quad (1)$$

$$Q_{1,j}=\mu_{Bi}(m)j=1, 2 \quad (2)$$

The product of the initial signal appears as the output of layer 2 including rule nodes, which is explained as:

$$W_i=\mu_{Ai}(m)\mu_{Bi}(m)j \quad (3)$$

Normalizing the weight function is applied by layer 3 (normalizing layer) as below:

$$w = \frac{w_i}{w_1 + w_2},\ i = 1, 2 \tag{4}$$

In layer 4, the output of prior layer multiplied with the function of Sugeno fuzzy rule is consequent nodes:

$$\overline{w}_i f_i = w_i(m_{p1} + n_{p1} + r_1) i = 1, 2 \tag{5}$$

The summation of all outputs in each rule beginning from the prior layer is calculated in layer 5 (output node).

$$Q^5 = \sum_{i=1}^{n} \overline{w}_i f_i \tag{6}$$

Layer 1 is fuzzy layer in which m is the input of nodes A_1, B_1 and also n is the input of nodes A_2, B_2. To divide the membership functions, labels A_1, B_1 and A_2, B_2 are utilized in the fuzzy theory. Layer 2 under the name of product layer contains two nodes. Labels w_1 and w_2 are outputs, namely the weight functions of the next layer.

3. Genetic Algorithm (GA)

GA is a randomly determined algorithm to find optimal solutions according to natural selection and genetics ideas. The efficient potency of GA for searching very large solution spaces can be defined as using probabilistic transition rules as an alternative of deterministic ones. GA covers three main stages: population initialization, GA operators and evaluation [70].

3-1. Population Initialization

Creation of a primary population, including primary nominee programs, is done randomly in GA. Each program in GA, namely a chromosome or string, is demonstrated by variable values which could be either binary-valued (binary coding) or real-valued numbers. In the current article, a binary coding scheme was employed by the GA model. The primary population should include strings adequate for all the optimization constraints.

3-2. GA Operators

The Darwinian main stage of reproduction and selection of the fittest and the genetic operations of crossover and mutation are employed to generate offspring computer solutions from the primary population. The reproduction includes choosing the solutions with the best fitness values and transferring these solutions to the novel generation. The cross-over operator imitates sexual reproduction involving two parents to generate offspring solutions. The various solutions are crossed over at random positions, i.e., arbitrary branches or sub-solutions are swapped among parents. The obtained offspring solutions are in various sizes and shapes compared to their original values and have dissimilar fitness values. Mutation is stimulated to randomly change the particular sections of the computer solutions [70].

3-3. Evaluation

In the evaluation process, the proposed solutions from the abovementioned procedures are assessed in the objective function of the optimization problem.

This procedure is reiterated to achieve the maximum allowable iterations or acceptable precision for the optimization issue [71].

In this regard, the mean square error (MSE) involving the estimated and actual values, computed as follows, is considered as the termination criteria in the optimization process:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(x_i^{exp} - x_i^{sim})^2 \tag{7}$$

where n is number of data points x^{sim} and x^{exp} demonstrate estimated and actual values, respectively.

4. Polynomial Regression (PR)

Ordinary least squares (OLS) regression is dominated by the following correlation:

$$\beta = (X'X)^{-1}X' \tag{8}$$

where β, X and Y are the vector of regression coefficient, the design matrix and the vector of responses at each data point, respectively. Multivariate polynomial regression (MPR) as a particular model of multivariate OLS regression explains the relationship of regressors with the response variable by the use of standard polynomial. Standard polynomial can be defined as the polynomial function, including every polynomial term conveyed by a multinomial expanded regressors (the degree is given). Also, the number of variables and polynomials of various degrees are inquired systematically to see which is the most appropriate fit. There is a limited number of standard polynomials investigated based on the degree of freedom which is proposed by a specific dataset. Note that the number of data points is always greater than the number of terms in the polynomial, which is in proportion with number of coefficients. MPR models produce results according to the effect of the estimator variable's combined interactions on the response. Nevertheless, all regression approaches have a restriction that explains the uncertainty about underlying mechanisms and only assures the relationships [72].

Ahmet Cecen [73,74] used a MATLAB code to develop the MPR model. In this regard, the order of MPR model was altered and subsequently we found the most accurate model. More details regarding the MPR model are reported in the literature [73].

RESULTS AND DISCUSSION

1. Data Attainment

To construct a precise and accurate model, using credible experimental data is very important [75,76]. For this purpose, 80 data points of H_2S solubility in [bmim][PF_6] were provided from the published literature [15,49]. Details of parameters used in this study are presented in Table 1. For each model, pressure and temperature are the input parameters, while the output parameter is CO_2 solubility in the mixed [bmim][PF_6] solvent.

2. MD Results

The adsorption isotherm task of Metropolis method was used in the calculations, which is based on the exchange step type, comprising creation of a new sorbate and deletion of an existing molecule in a given temperature and a range of fugacity. In our work

Table 1. Details of the parameters used in the current article

Variable	Status	Min	Max
Temperature (K)	Input	298.15	403.15
Pressure (kPa)	Input	0.8	9630
H_2S solubility (mole H_2S/mole [bmim][PF_6])	Output	0.016	0.875

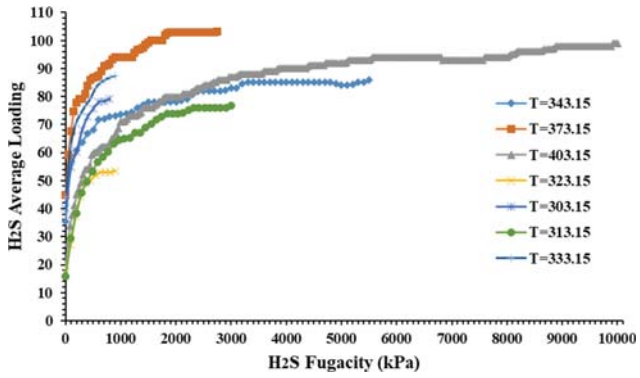

Fig. 5. Average loading of H_2S in [bmim][PF_6] for different isotherms in terms of gas fugacity.

the quality of calculation was set to fine in sorption module, and the adsorption isotherm was obtained at various temperatures. The force field used for both H_2S molecules and IL was COMPASS II, while the atom-based method was chosen for electrostatics as well as van der Waals interactions. Simulations were carried out for 3×10^6 trial configurations, which consisted of 1×10^6 cycles for the equilibration and 2×10^6 cycles for the production step.

The solubility of H_2S into [bmim][PF_6] IL for several temperatures was obtained using MC simulation. Pressure range for each isotherm was chosen according to experiment pressure. Fig. 5 illustrates the adsorption isotherms at various temperatures. The range of fugacity for each temperature corresponds to the experimental fugacity. In some temperatures, the available experimental fugacity reached 10,000 Kilo Pascal (kPa), whereas for some others it was below 1,000 KPa. The adsorption isotherm of H_2S solubility in [bmim][PF_6] at all temperatures studied here is of type I adsorption. The obtained graphs depict monolayer adsorption, which can be easily explained using the Langmuir adsorption isotherm.

3. PR

To approximate the parameters of models, nonlinear regression method considered as a general approach can be employed, even if the probable model cannot be linearized. Input variables of H_2S solubility function (P, T) are expressed as follows: pressure (KPa) and temperature (K). The results of nonlinear regression modeling are described in the below formula:

$$\begin{aligned} H_2S\ \text{Solubility} = {} & +-8.0109e\text{-}7P^2+1.7792e\text{-}11P^3+2.3759e\text{-}5TP \\ & +3.6588e\text{-}9TP^2+-4.1371e\text{-}14TP^3+3.3234e\text{-}6T^2 \\ & +-1.1662e\text{-}7T^2P+-4.2374e\text{-}12T^2P^2 \\ & +-1.8475e\text{-}8T^3+1.4521e\text{-}10T^3P \\ & +01+2.54e\text{-}11T^4+-3.8227e\text{-}17P^4 \end{aligned} \quad (9)$$

At this point, as presented in Eq. (9), the interaction terms of the 4-order polynomial model were close-fitting to the training set where the remaining variables were used as the estimators in order to predict the output.

4. ANFIS Modeling

The primary FIS was created by using the genfis3 function of MATLAB software. Regardless the importance of this necessity, estimating the solubility of H_2S in [bmim][PF_6] according to operating conditions was the purpose of this study. The pressure and temperature were set as inputs. Since, the H_2S solubility was considered as output.

Calculating the type and the number of the membership functions and also the quantity of iterations (epoch number) are the first steps to start the GA-ANFIS approach. The determination of these values is very important, as that sets the performance of the model.

Processing experimental data into specific patterns has to be mentioned first before the ANFIS can examine and the mapping could learn. The two common steps can be defined as follows: forming a pattern vector is the first one; pattern constitution including an input condition vector and the corresponding target vector is considered as the second step. Additionally, the essential issue, which is considerable, is the range of input and output data. It becomes more engrossing especially in different variables of operating ranges. Selecting the input variable is the initial priority in modeling of nonlinear systems such as ANFIS. Hence, the data should be divided into the train and test datasets, where the 80 investigational data are divided into training (60) and test (20) data sets randomly.

Commonly, modeling a system with FL does not yield a precise outcome as a result of the structural bugs in the algorithm. Hence, other techniques like ANN are employed for pattern recognition. Coupling the power of FL and ANN creates a network, namely ANFIS. This network employs membership functions (MFs) and fuzzy if-then rules to create a fuzzy inference system (FIS). The MFs are adjusted by ANNs [77]. GA is used to optimize ANFIS structure.

Estimating the number of necessary MFs according to absence of explicit formula is inapplicable, as mentioned in literature. As recommended by many studies, the Gaussian was used as membership functions in the present study (e.g., [78-80]). The GA-ANFIS approach was coded with MATLAB software package. To estimate the H_2S solubility in [bmim][PF_6] GA-ANFIS approach was developed. Table 2 shows the main GA parameters. These tabled parameters, which are the optimum values for this case study, were specified by trial and error procedure. The process of GA optimization of ANFIS structure is shown in Fig. 6 On the other side, Fig. 7 shows the structure of the GA-ANFIS and consistent membership functions, respectively. It demonstrates trained MF variables for input variables, i.e., pressure and temperature for each cluster. The GA-ANFIS has ten clusters.

Table 2. Properties of the GA-ANFIS model for estimating the solubility of H_2S

Variable	Value
Population number	60
Size of generation (iterations)	1000
Mutation %	0.9
Mutation rate	0.05
Crossover %	0.65
Lower bound of variables	−10
Upper bound of variables	10
No. of fuzzy rules	10

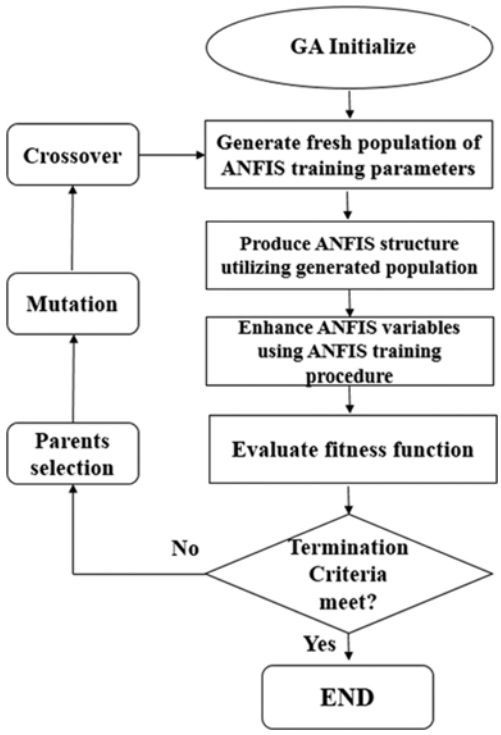

Fig. 6. GA optimization procedure.

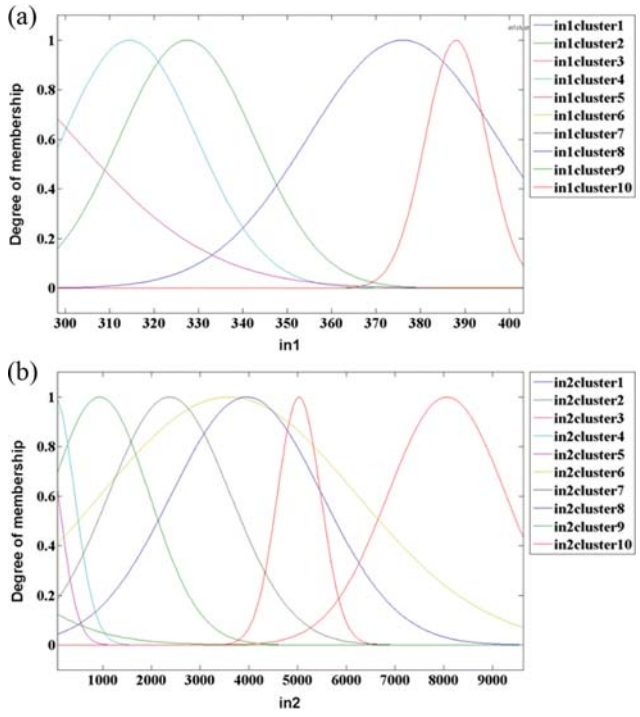

Fig. 7. The MFs for (a) temperature and (b) pressure for the GA-ANFIS.

For measuring the model's performance, the error test is commonly employed as a genuine determination. The best model regarded as the owner of minimal test error is connoted by the results.

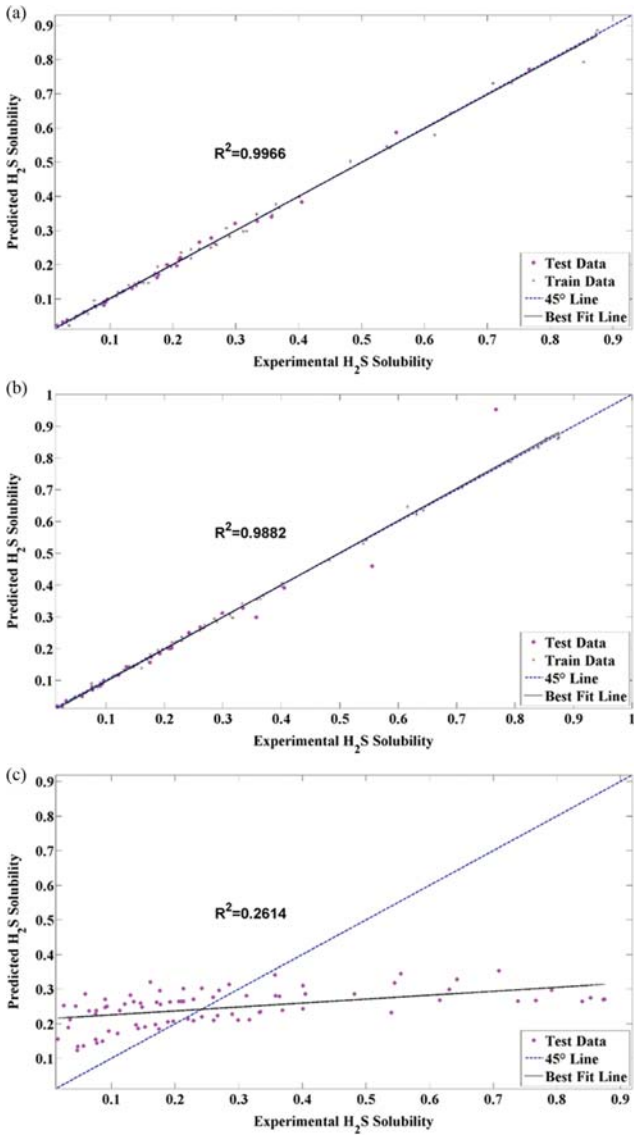

Fig. 8. Predicted H_2S solubility data vs. experimental data for the models: (a) GA-ANFIS, (b) PR, (c) MS.

5. Precision of the Recommended Model and Evaluation

By the use of the mean squared errors (MSEs), average absolute relative error (%AARD), standard deviation (STD) and the coefficient of determination (R^2) between the experimental and estimated data, the performance of the proposed models was measured. Eqs. (11)-(13) show the formulation of the % AARD, STD and R^2 as follows [81-84]:

$$R^2 = 1 - \frac{\Sigma_{i=1}^{n}(x_i^{sim} - x_i^{exp})^2}{\Sigma_{i=1}^{n}(x_i^{sim} - x^{m})^2},\ x_m = \frac{\Sigma_{i=1}^{n} x_i^{exp}}{n} \tag{10}$$

$$\%\ AARD = \frac{100}{n}\Sigma_{i=1}^{n}\left|\frac{x_i^{sim} - x_i^{exp}}{x_i^{exp}}\right| \tag{11}$$

$$STD = \Sigma_{i=1}^{n}\left(\frac{(x_i^{sim} - x_m)^2}{n}\right)^{0.5} \tag{12}$$

The cross plot of estimated values versus the target values is represented for graphical methods in Fig. 8(a), 8(b) and 8(c) under the name of GA-ANFIS, PR and MS, respectively. In these figures, the vertical and horizontal axes are, respectively estimated and investigational values of H_2S solubility in ILs. If more compact data are allocated to y=x, the estimated value will be more precise. As it is

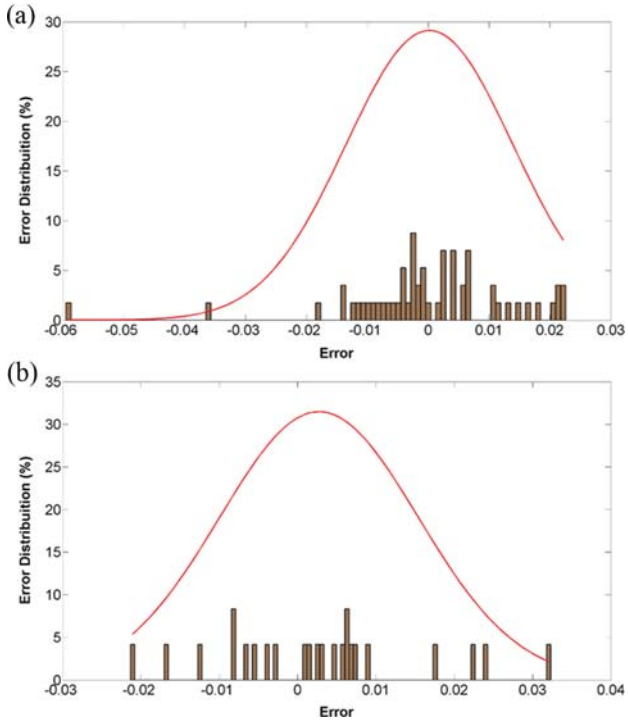

Fig. 9. Histogram of errors for the H_2S solubility estimated by GA-ANFIS models: (a) training and (b) testing data.

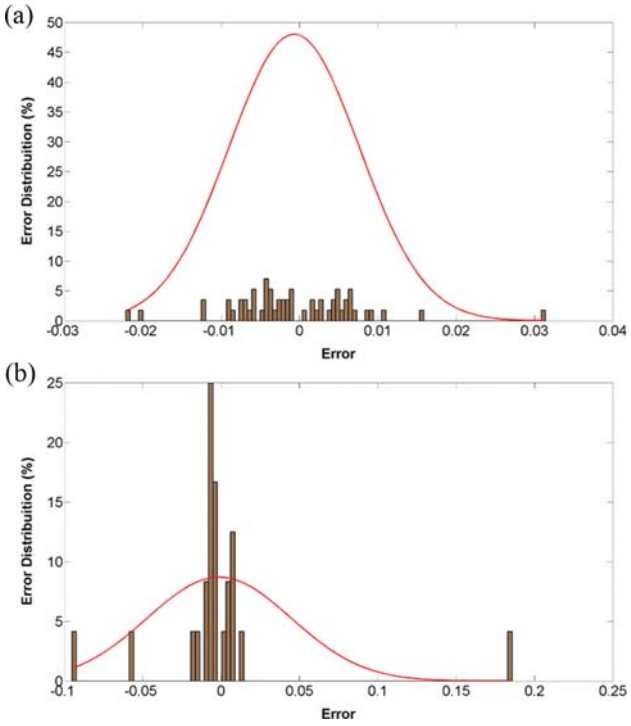

Fig. 10. Histogram of errors for the H_2S solubility estimated by PR models: (a) training and (b) testing data.

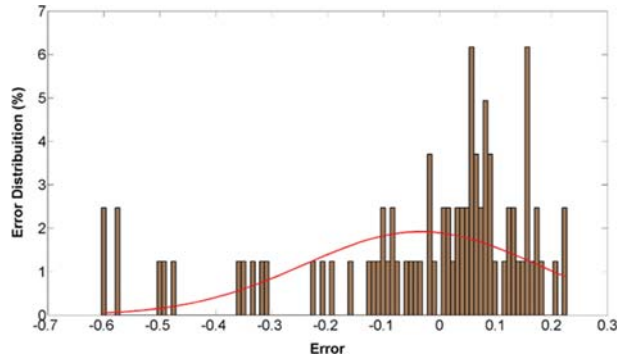

Fig. 11. Histogram of errors for the H_2S solubility estimated by MS models.

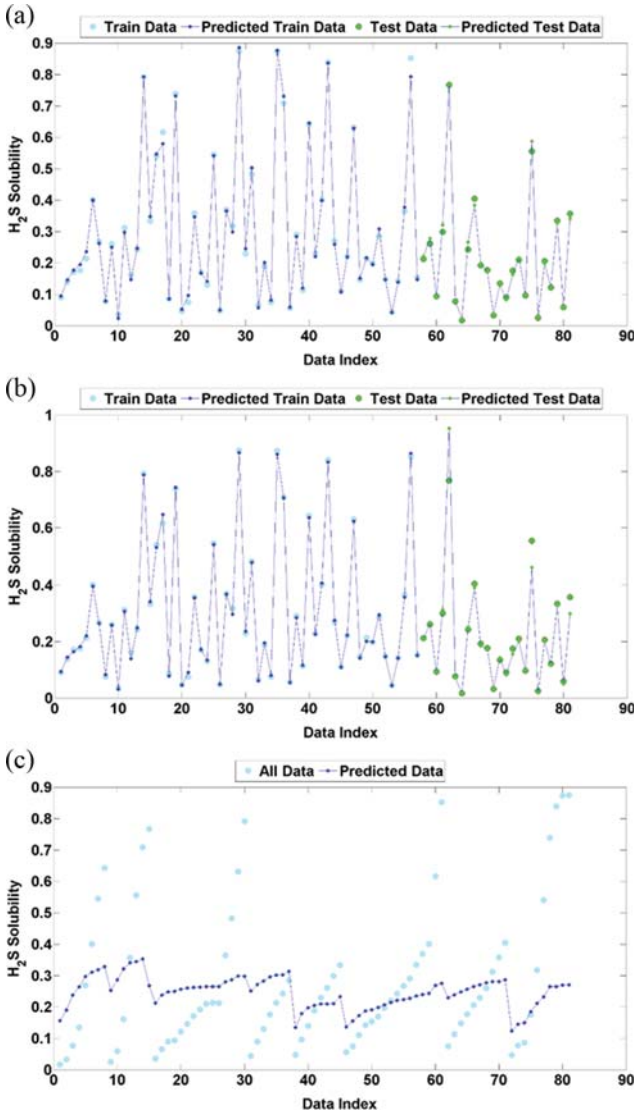

Fig. 12. Trend plot of actual and estimated data vs. sequence of data points for: (a) GA-ANFIS, (b) PR, (c) MS.

Table 3. Comparison of the performance of the models

Model analysis	GA-ANFIS			PR			MS
	Training	Testing	All	Training	Testing	All	All
R^2	0.9969	0.9951	0.9966	0.9988	0.9524	0.9882	0.2614
MSE	0.0002	0.0002	0.0002	0.0001	0.0020	0.0006	0.0438
STD	0.2395	0.1749	0.2262	0.2412	0.1930	0.2320	0.0609
%AARD	4.8338	7.6753	5.6758	3.1568	8.5516	4.7553	101.8296

clear, the values predicted by all the predictors and experimental data are consistent in the approach. The R^2 of GA-ANFIS, PR and MD is 0.9986, 0.9882 and 0.2614, respectively. As shown, GA-ANFIS followed by PR has the highest R^2 or accuracy. The histogram of errors is represented in Figs. 9-11. It can be observed that none of the models follows the standard normal distribution. To acquire an obvious comprehension regarding the accuracy of the proposed model, the trend plots of the employed model for the prediction performance are exhibited in Fig. 12. Fig. 12 reveals that the prediction of the GA-ANFIS and PR models follows the trend of actual data with a satisfactory level of preciseness, as can be seen from the good overlap involving the actual data and model estimations.

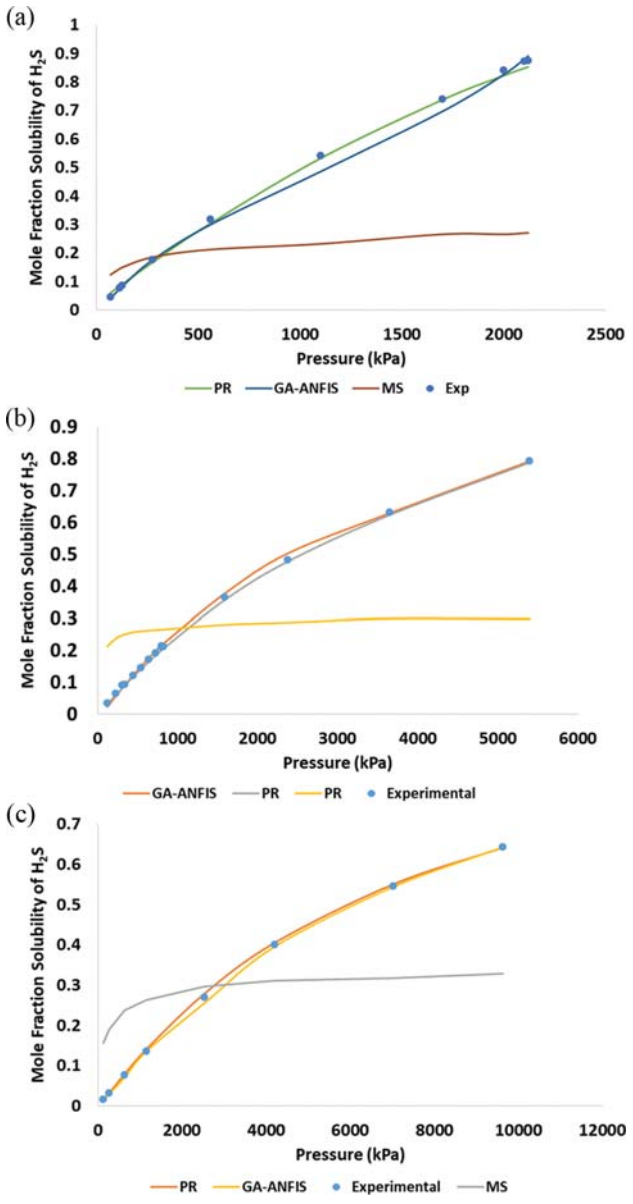

Fig. 13. Comparison between estimated H_2S solubility data and experimental values: (a) T=298.15 K, (b) T=343.15 K and (c) T=403.15 K.

The values of MSE, %AARD, STD and, R^2 of data set are listed in Table 3 because of statistical criteria employed for the GA-ANFIS, PR and MS models. The output values of GA-ANFIS and PR are in a good accordance with actual values. The obtained values of %AARD are 7.68 for GA-ANFIS, 8.55 for PR and 101.83 for MS. The obtained values of R^2 are 0.9951 for GA-ANFIS 0.9524 for PR and 0.2614 for MS. In addition, the obtained values of MSE are 0.002 for GA-ANFIS 0.002 for PR and 0.0438 for MS. It is obvious that the GA-ANFIS model performs better than PR and MD models for H_2S solubility and represents better results. The offered hybrid artificial intelligence model can estimate more precise data for nonlinear systems. Nevertheless, the accuracy of simple yet accurate PR model is very high, and this model can be easily used for industrial purposes.

To evaluate the proposed models, GA-ANFIS, PR models and MS results were compared with experimental data at different pressure and temperatures. As shown in Fig. 13, three different diagrams were plotted at 298.15, 343.15 and 403.15 K. All the diagrams present the mole fraction of H_2S solubility versus pressure (kPa). As can be seen in all diagrams, the H_2S solubility increases with pressure, which means that all the suggested models follow the same trend. Note that the deviation of MS results is considerable, although the obtained results from other models and experimental data are consistent in the approach. It must be mentioned that PR and GA-ANFIS models estimate H_2S solubility only on the basis of the trained data in the training stage of modeling process and they differ from the thermodynamics models which their predictability is different at lower or higher pressure and temperatures.

CONCLUSION

GA-ANFIS, MD and PR, which are three different methods, were employed to approximate the solubility of H_2S in [bmim][PF_6] IL magnificently. According to graphical and statistical representation methods, actual data and estimated values are consistent with each other. Furthermore, the obtained results of PR model and experimental (actual) data are consistent in approach because of

higher R^2 equal to 0.9967 and lower values of %AARD 2.0812. In addition, the concordance of MD simulation results and experimental data is less consistent than other two methods because of the nature of this method and originality of producing data based on the physics and chemistry of the material. The predictability of this method can be improved by improvement of the force field used. This research can be considered as a perspective on H_2S removal and also in gas processes where accurate and precise data for solubility of H_2S is necessary.

ABBREVIATIONS

AA-FF : all-atom force field
AARD : average absolute relative deviation
AARD : average absolute relative error
ANN : artificial neural networks
EOS : equation of states
ESP : electro-static potential
GA-ANFIS: genetic algorithm-Adaptive neuro fuzzy inference system
H2S : hydrogen sulfide
MC : monte carlo
MD : molecular dynamics
MPR : multivariate polynomial regression
MSE : mean squared errors
NTRL : non-random two-liquid
PR : polynomial regression
PR : Peng-Robinson
PSO : particle swarm optimization
R2 : coefficient of determination
SAFT : statistical associating fluid theory
SRK : Soave-Redlich-Kwong
STD : standard deviation
SVM : support vector machines
UA-FF : united-atom force field

REFERENCES

1. N. Ai, J. Chen and W. Fei, *J. Chem. Eng. Data*, **50**, 492 (2005).
2. A. Dashti, H. R. Harami and M. Rezakazemi, *Int. J. Hydrogen Energy*, **43**, 6614 (2018).
3. M. Rezakazemi, A. Dashti, M. Asghari and S. Shirazian, *Int. J. Hydrogen Energy*, **42**, 15211 (2017).
4. A. Dashti and M. Asghari, *ChemBioEng Rev.*, **2**, 54 (2015).
5. J. L. Anthony, E. J. Maginn and J. F. Brennecke, Gas solubilities in 1-n-butyl-3-methylimidazolium hexafluorophosphate, ACS Symposium Series (2002).
6. E. D. Bates, R. D. Mayton, I. Ntai and J. H. Davis, *J. Am. Chem. Soc.*, **124**, 926 (2002).
7. T. Welton, *Chem. Rev.*, **99**, 2071 (1999).
8. J. F. Brennecke and E. J. Maginn, *AIChE J.*, **47**, 2384 (2001).
9. Q. Yang and D. D. Dionysiou, *J. Photochem. Photobiol. A: Chem.*, **165**, 229 (2004).
10. K. Seddon, *Kinet. Catal.*, **37**, 693 (1995).
11. C. Lagrost, D. Carrie, M. Vaultier and P. Hapiot, *J. Phys. Chem. A*, **107**, 745 (2003).
12. A. Shariati and C. J. Peters, *J. Supercrit. Fluids*, **34**, 171 (2005).
13. A. Shariati, K. Gutkowski and C. J. Peters, *AIChE J.*, **51**, 1532 (2005).
14. H. Zhao, S. Xia and P. Ma, *J. Chem. Technol. Biotechnol.*, **80**, 1089 (2005).
15. A. H. Jalili, M. Rahmati-Rostami, C. Ghotbi, M. Hosseini-Jenab and A. N. Ahmadi, *Chem. Eng. Data*, **54**, 1844 (2009).
16. M. Shokouhi, M. Adibi, A. H. Jalili, M. Hosseini-Jenab and A. Mehdizadeh, *J. Chem. Eng. Data*, **55**, 1663 (2009).
17. A. H. Jalili, A. Mehdizadeh, M. Shokouhi, A. N. Ahmadi, M. Hosseini-Jenab and F. Fateminassab, *J. Chem. Thermodyn.*, **42**, 1298 (2010).
18. A. Shariati, S.-S. Ashrafmansouri, M. H. Osbuei and B. Hooshdaran, *Korean J. Chem. Eng.*, **30**, 187 (2013).
19. X. Ji and H. Adidharma, *Fluid Phase Equilib.*, **293**, 141 (2010).
20. A. Eslamimanesh, F. Gharagheizi, A. H. Mohammadi and D. Richon, *Chem. Eng. Sci.*, **66**, 3039 (2011).
21. J. S. Torrecilla, J. Palomar, J. García, E. Rojo and F. Rodríguez, *Chemometrics Intellig. Lab. Syst.*, **93**, 149 (2008).
22. P. F. Arce, P. A. Robles, T. A. Graber and M. Aznar, *Fluid Phase Equilib.*, **295**, 9 (2010).
23. A. Perez-Salado Kamps, D. Tuma, J. Xia and G. Maurer, *J. Chem. Eng. Data*, **48**, 746 (2003).
24. S. K. Shukla, S. G. Khokarale, T. Q. Bui and J.-P. Mikkola, *Front. Mater.*, **6** (2019).
25. S. H. H. N. Ghazani, A. Baghban, A. H. Mohammadi and S. Habibzadeh, *J. Supercrit. Fluids*, **133**, 455 (2018).
26. M. Mesbah, S. Shahsavari, E. Soroush, N. Rahaei and M. Rezakazemi, *J CO2 Util.*, **25**, 99 (2018).
27. D.-Y. Peng and D. B. Robinson, *Ind. Eng. Chem. Res.*, **15**, 59 (1976).
28. G. Soave, *Chem. Eng. Sci.*, **27**, 1197 (1972).
29. L. F. Vega, O. Vilaseca, F. Llovell and J. S. Andreu, *Fluid Phase Equilib.*, **294**, 15 (2010).
30. H. Renon and J. M. Prausnitz, *AIChE J.*, **14**, 135 (1968).
31. G. M. Wilson, *J. Am. Chem. Soc.*, **86**, 127 (1964).
32. D. S. Abrams and J. M. Prausnitz, *AIChE J.*, **21**, 116 (1975).
33. F. Aa, J. Gmehling and P. Rasmussen, Vapor-Liquid Equilibria Using UNIFAC: A Group-Contribution Method, AmsterdamB Elsevier, Amsterdam (1977).
34. J. Gross and G. Sadowski, *Ind. Eng. Chem. Res.*, **40**, 1244 (2001).
35. E. K. Karakatsani and I. G. Economou, *J. Phys. Chem. B*, **110**, 9252 (2006).
36. S.-S. Ashrafmansouri and S. Raeissi, *J. Supercrit. Fluids*, **63**, 81 (2012).
37. M. Safamirzaei and H. Modarress, *Thermochim. Acta*, **545**, 125 (2012).
38. M. Safamirzaei and H. Modarress, *Fluid Phase Equilib.*, **332**, 165 (2012).
39. E. J. Maginn, *Acc. Chem. Res.*, **40**, 1200 (2007).
40. E. J. Maginn, *J. Phys.: Condens. Matter.*, **21**, 373101 (2009).
41. P. J. G. Huttenhuis, N. Agrawal, J. Hogendoorn and G. Versteeg, *J. Petrol. Sci. Eng.*, **55**, 122 (2007).
42. C. Hanke, A. Johansson, J. Harper and R. Lynden-Bell, *Chem. Phys. Lett.*, **374**, 85 (2003).
43. J. K. Shah, J. F. Brennecke and E. J. Maginn, *Green Chem.*, **4**, 112 (2002).
44. T. I. Morrow and E. J. Maginn, *J. Phys. Chem. B*, **106**, 12807 (2002).

45. J. K. Shah and E. J. Maginn, *Fluid Phase Equilib.*, **222**, 195 (2004).
46. B. Widom, *J. Chem. Phys.*, **39**, 2808 (1963).
47. C. S. Pomelli, C. Chiappe, A. Vidis, G. Laurenczy and P. J. Dyson, *J. Phys. Chem. B*, **111**, 13014 (2007).
48. A. H. Jalili, M. Safavi, C. Ghotbi, A. Mehdizadeh, M. Hosseini-Jenab and V. Taghikhani, *J. Phys. Chem. B*, **116**, 2758 (2012).
49. F.-Y. Jou and A. E. Mather, *Int. J. Thermophys.*, **28**, 490 (2007).
50. J. L. Anthony, J. L. Anderson, E. J. Maginn and J. F. Brennecke, *J. Phys. Chem. B*, **109**, 6366 (2005).
51. C. Cadena, J. L. Anthony, J. K. Shah, T. I. Morrow, J. F. Brennecke and E. J. Maginn, *J. Am. Chem. Soc.*, **126**, 5300 (2004).
52. J. Jacquemin, P. Husson, V. Majer and M. F. C. Gomes, *Fluid Phase Equilib.*, **240**, 87 (2006).
53. M. A. Ahmadi, R. Haghbakhsh, R. Soleimani and M. B. Bajestani, *J. Supercrit. Fluids*, **92**, 60 (2014).
54. A. Shafiei, M. A. Ahmadi, S. H. Zaheri, A. Baghban, A. Amirfakhrian and R. Soleimani, *J. Supercrit. Fluids*, **95**, 525 (2014).
55. M.-A. Ahmadi, B. Pouladi, Y. Javvi, S. Alfkhani and R. Soleimani, *J. Supercrit. Fluids*, **97**, 81 (2015).
56. A. Baghban, M. A. Ahmadi and B. H. Shahraki, *J. Supercrit. Fluids*, **98**, 50 (2015).
57. L. Xia, J. Wang, S. Liu, Z. Li and H. Pan, *Processes*, **7**, 258 (2019).
58. M. J. Muldoon, S. N. Aki, J. L. Anderson, J. K. Dixon and J. F. Brennecke, *J. Phys. Chem. B*, **111**, 9001 (2007).
59. J. Liang, R. Zhang, Q. Zhao, J. Dong, B. Wang and J. Li, *Comput. Theor. Chem.*, **980**, 1 (2012).
60. R. A. Sarmiento-Perez, L. M. Rodriguez-Albelo, A. Gomez, M. Autie-Perez, D. W. Lewis and A. R. Ruiz-Salvador, *Micropor. Mesopor. Mater.*, **163**, 186 (2012).
61. S. Velioğlu, M. G. Ahunbay and S. B. Tantekin-Ersolmaz, *J. Membr. Sci.*, **417**, 217 (2012).
62. C. Liu, Y. Dang, Y. Zhou, J. Liu, Y. Sun, W. Su and L. Zhou, *Adsorpt.*, **18**, 321 (2012).
63. M. Li, X. Huang and Z. Kang, *J. Appl. Phys.*, **118**, 084303 (2015).
64. Z. Lei, J. Jiang, G. Zhu, P. Cui, Q. Ling and Z. Zhao, *Energy Fuels*, **30**, 1287 (2016).
65. A. D. Becke, *J. Chem. Phys.*, **98**, 5648 (1993).
66. C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, **37**, 785 (1988).
67. B. Derecskei and A. Derecskei-Kovacs, *Mol. Simulat.*, **34**, 1167 (2008).
68. J.-S. R. Jang, *IEEE. T. Syst. Man. CY.*, **23**, 665 (1993).
69. T. Mohammadi and A. Esmaeelifar, *Desalination*, **166**, 329 (2004).
70. X. Chen and N. Wang, *Chem. Eng. J.*, **150**, 527 (2009).
71. L. Davis, Handbook of genetic algorithms, Van Nostrand Reinhold, New York (1991).
72. K. P. Singh, S. Gupta, A. Kumar and S. P. Shukla, *Sci. Total Environ.*, **426**, 244 (2012).
73. A. Cecen, Calculation, utilization, and inference of spatial statistics in practical spatio-temporal data, Georgia Institute of Technology (2017).
74. A. Agrawal, P. D. Deshpande, A. Cecen, G. P. Basavarsu, A. N. Choudhary and S. R. Kalidindi, *Integr. Mater. Manuf. Innov.*, **3**, 8 (2014).
75. A. Shokrollahi, A. Tatar and H. Safari, *J. Taiwan Inst. Chem. E.*, **55**, 17 (2015).
76. H. Darvish, M. Nouri-Taleghani, A. Shokrollahi and A. Tatar, *J. Afr. Earth. Sci.*, 111, **409** (2015).
77. J.-S. R. Jang, C.-T. Sun and E. Mizutani, *IEEE Trans Automat Contr.*, **42**, 1482 (1997).
78. M. Rezakazemi, A. Ghafarinazari, S. Shirazian and A. Khoshsima, *Polym. Eng. Sci.*, **53**, 1272 (2013).
79. L. Yingjie and W. Baoshu, *J. Sys. Eng. Electron.*, **16**, 583 (2005).
80. A. Dashti, M. Raji, A. Azarafza, A. Baghban, A. H.Mohammadi and M. Asghari, *J. Env. Manage.*, **224**, 58 (2015).
81. A. Dashti, M. Raji, A. Razmi, N. Rezaei, S. Zendehboudi and M. Asghari, *Chem. Eng. Res. Des.*, **144**, 405 (2019).
82. M. Raji, A. Dashti, P. Amani and A. H. Mohammadi, *J. Mol. Liq.*, **283**, 804 (2019).
83. A. Dashti, H. R. Harami, M. Rezakazemi and S. Shirazian, *J. Mol. Liq.*, **271**, 661 (2018).
84. M. Rezakazemi, A. Azarafza, A. Dashti and S. Shirazian, *Int. J. Hydrogen Energy*, **43**, 17283 (2018).